

Amendments to the Specification

Please replace the paragraph beginning at page 7, line 12, with the following rewritten paragraph:

Query processing component 240 receives information requests from a requestor 110. Query processing component 240 may analyze the information requests to understand better what information the requestor 110 seeks. Particularly salient to the present invention, query processing component 240 may determine equivalent descriptions for the information need sought by requestor 110. Query processing component 240 may also perform functions such as detecting and correcting misspellings, ~~transform~~ transforming the information request into a format that is more likely to produce the desired information, when processed by the remainder of search engine 205, etc.

Please replace the paragraph beginning at page 8, line 10, with the following rewritten paragraph:

Output component 270 receives an ordered list of items from ordering component 260, formats that list ~~into~~ in a manner suitable for presenting to requestor 110, and provides the ordered list to requestor 110.

Please replace the paragraph beginning at page 13, line 15, with the following rewritten paragraph:

Fig. 7 is a flow diagram for identifying equivalent descriptions based on anchor text information, consistent

with the present invention. For purposes of illustration, this process will be described in conjunction with the document system shown in Fig. 8, which shows three documents: A, B, C and D. Document A contains anchor text "palo alto cars", which is hyperlinked to document E D, and anchor text "palo alto inns", which is hyperlinked to document D C. Document B contains anchor text "san francisco inns", which is also hyperlinked to document D C. Document E D presumably contains information about palo alto cars, while document D C presumably contains information about palo alto inns and san francisco inns.

Please replace the paragraph beginning at page 14, line 6, with the following rewritten paragraph:

The anchor text units are then organized by the document being pointed to. (Stage 720). Accordingly, "palo alto cars" would be in one set, since it points to document E D; and "palo alto inns" and "san francisco inns" would be in a second set, since they each point to document D C.

Please replace the paragraph beginning at page 14, line 15, with the following rewritten paragraph:

A score is calculated for each candidate pair (stage 750) and the terms of the candidate pair ~~is~~ are treated as synonyms if the score exceeds a threshold (stage 760). The score for each candidate pair may be calculated in a variety of ways, similar to those described above in reference to Fig. 5. For example, if the score may be

calculated as a ratio of the number of times the transform occurs divided by the total number of times the first half of the candidate pair occurs in the entire collection of anchor text units. Furthermore, the threshold may be set depending on the desired confidence level, also similar to that described above in reference to Fig. 5.